

Optimizing Motor Intention Detection with Deep Learning: Towards Management of Intraoperative Awareness

Oleksii Avilov, Sébastien Rimbert, Anton Popov, Laurent Bougrain

▶ **To cite this version:**

Oleksii Avilov, Sébastien Rimbert, Anton Popov, Laurent Bougrain. Optimizing Motor Intention Detection with Deep Learning: Towards Management of Intraoperative Awareness. IEEE Transactions on Biomedical Engineering, 2021, pp.10. 10.1109/TBME.2021.3064794 . hal-03166495

HAL Id: hal-03166495
https://inria.hal.science/hal-03166495

Submitted on 11 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimizing Motor Intention Detection with Deep Learning: Towards Management of Intraoperative Awareness

Oleksii Avilov, *Member, IEEE*, Sébastien Rimbert, Anton Popov, *Senior Member, IEEE*, and Laurent Bougrain

***Abstract*— Objective: This article shows the interest in deep learning techniques to detect motor imagery (MI) from raw electroencephalographic (EEG) signals when a functional electrical stimulation is added or not. Impacts of electrode montages and bandwidth are also reported. The perspective of this work is to improve the detection of intraoperative awareness during general anesthesia.**

Methods: Various architectures of EEGNet were investigated to optimize MI detection. They have been compared to the state-of-the-art classifiers in Brain-Computer Interfaces (based on Riemannian geometry, linear discriminant analysis), and other deep learning architectures (deep convolution network, shallow convolutional network). EEG data were measured from 22 participants performing motor imagery with and without median nerve stimulation.

Results: The proposed architecture of EEGNet reaches the best classification accuracy (83.2%) and false-positive rate (FPR 19.0%) for a setup with only six electrodes over the motor cortex and frontal lobe and for an extended 4-38 Hz EEG frequency range while the subject is being stimulated via a median nerve. Configurations with a larger number of electrodes result in higher accuracy (94.5%) and FPR (6.1%) for 128 electrodes (and respectively 88.0% and 12.9% for 13 electrodes).

Conclusion: The present work demonstrates that using an extended EEG frequency band and a modified EEGNet deep neural network increases the accuracy of MI detection when used with as few as 6 electrodes which include frontal channels.

Significance: The proposed method contributes to the development of Brain-Computer Interface systems based on MI detection from EEG.

***Index Terms*— Brain-computer interface (BCI), deep learning, electroencephalogram (EEG), intraoperative awareness during general anesthesia, machine learning, median nerve stimulation, motor imagery.**

This paragraph of the first footnote will contain the date on which you submitted your paper for review. O.A. was supported by scholarship from the French Embassy to Ukraine while working on this topic at the NEUROSYS team at LORIA (Université de Lorraine/CNRS/Inria), Nancy, France. Experiments presented in this paper were carried out using the Grid'5000 testbed, supported by a scientific interest group hosted by Inria and including CNRS, RENATER and several Universities as well as other organizations (https://www.grid5000.fr).

O. A., S. R. and L. B. are with Université de Lorraine, CNRS, Inria, LORIA, F-54000, Nancy, France

O. A. and A. P. are with Electronic Engineering Department, Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, 03056, Ukraine (e-mail: oleksii.avilov@loria.fr).

Acronyms: AAGA: accidental awareness during general anesthesia, BCI: brain-computer interface, CSP: common spatial pattern, CNN : convolutional neural networks, EEG: electroencephalography, ERD/ERS: event-related desynchronization/event-related synchronization, ERSP: event-related spectral perturbation, FBCSP: filter bank common spatial pattern, FPR : false-positive rate, LDA: linear discriminant analysis, LR: logistic regression, MDRM : minimal distance to the riemannian mean, MI: Motor imagery, MNS: median nerve stimulation, PTSD: post-traumatic stress disorder, TS: tangent space.

I. Introduction

HUNDREDS of millions of general anesthesia are performed each year on patients all over the world [1]. Among these patients, 0.2-1.3% suffer from Accidental Awareness during General Anesthesia (AAGA) [2], i.e., an unexpected awakening during a surgical procedure under general anesthesia. It can lead to physical suffering or psychological damage called post-traumatic stress disorder (PTSD) [3]. People with explicit memories of PTSD may have long-term postoperative effects including anxiety, irritability, fear of death, sleep disturbances, and depressive syndromes [4].

Prevention of AAGA during surgery is a challenge because technologies used by anesthesiologists are not able to predict all patients' recovery. Currently, there are two ways to monitor the depth of general anesthesia: (i) by observing the clinical features (i.e., heart rate, blood pressure, movements, sweating), which may be objective or subjective in nature, (ii) or by recording a portion of the electroencephalographic (EEG) signal at the frontal cortex (i.e., *Bispectral Index*, *Patient State Index*, *Entropy*) [5]. Unfortunately, both clinical signs and new brain indices (based on frontal lobe activity) are not satisfactory enough to successfully assess the depth of general anesthesia and prevent AAGA [6], [7].

In previous works [8]–[10], the concept of an innovative passive brain-computer interface (BCI) based on detection of movement intentions to prevent AAGA was proposed. Indeed, during AAGA patients typically try to move to alert the medical staff, but discover that they are unable to do it [11]. The proposed solution incorporates a Median Nerve Stimulation (MNS) causing specific modulations of activity in the sensorimotor cortex [12], which are confirmed to be altered by an intention of movement [10], [13]. Indeed, during a movement or motor imagery, a minimal power level is

maintained in both alpha and beta bands [14], called event-related desynchronization (ERD). Then, 300–500 ms after the end of the task, there is an increase of power referred to as an event-related-synchronization (ERS) in the beta band, also known as a post-movement beta rebound, lasting 1 s [12]. Concurrently, in the alpha band, the power returns to a baseline after several seconds. More interestingly, a MNS induces an ERD during the stimulation while an ERS appears after the stimulation [12], [13], [15] but the ERS is strongly abolished if a MNS is performed during an intention of movement [9], [10], [13]. This interesting process could make the detection of AAGA with a passive BCI possible. Indeed, we can imagine a routine system where the patient would be stimulated at the median nerve. Moreover, to find out if the patient has an intention to move, the analysis of ERD and ERS modulations of the motor cortex would be used [10]. Paradigms, classifiers, frequency bands and electrode layouts studied in this article are promising to design a BCI that could help to detect AAGA.

In our previous studies [9], [10], it was already shown that a MI + MNS vs. MNS classification has better accuracy than a MI vs. Rest classification and could be used as a trigger in a system-paced BCI. The best classification performance was reached by a logistic regression (LR) based on a tangent space (TS) projection of covariance matrices (TS+LR) which obtained 70.5% of accuracy for a MI vs. resting-state task comparable to 82.03% for a MI + MNS vs. MNS task. Although the use of median nerve stimulation is promising, it still requires more detailed investigations, particularly concerning which frequency bands can be used to maximize discrimination between the two classes.

Other challenges remain for this kind of BCI to be used in clinical practice. Firstly, the accuracy obtained for a MI versus Rest classification in the BCI field in general remains low and should be improved to create a reliable device which can be used in hospitals. Typically, common spatial pattern (CSP) and filter bank common spatial pattern (FBCSP) algorithms are still the most popular methods for feature extraction [16] from EEG signal and showed both advantages (i.e., high accuracy, flexibility, simplicity) and limitations (i.e., frequency band selection). Recently, other algorithms based on Riemannian geometry (such as tangent space (TS) algorithm [10], [17] and minimum distance to Riemannian mean (MDRM) algorithm [17], [18]) showed promising results for MI detection. So, we will compare all of these methods with deep learning techniques in this study. Indeed, whereas deep neural networks show better performances compared to standard machine learning methods in almost all application domains, analysis of EEG signals especially in BCI domain remains an area where the use of deep learning techniques is still limited [19], [20]. The main reason is the small amount of data available for training and low signal-to-noise ratio of the EEG signals. The potential of deep learning architectures which can deal with raw EEG signals especially for motor imagery based tasks is rather unexplored and is active research field [20]. There are several deep learning architectures that already showed competitive results for MI vs. Rest task in BCI domain [20]–[22]. At the same time, we already showed potential of convolutional neural networks to classify MI + MNS vs. MNS task on the smaller dataset [23]. Also, one more challenge is to decrease number of EEG electrodes. Since applying EEG electrodes takes time and increase the preparation time it is often not possible to install 64 or 128 electrodes before the surgery. The solution might be to select the electrode setup which will show acceptable classification accuracy with minimum number of electrodes.

The objective of this study is to better investigate the ability of deep learning techniques with variants of convolutional neural networks (CNN) to extract features from raw EEG signals applied to the problem of detection of intraoperative awareness during general anesthesia especially when a MNS is applied. Our goal is not only to increase accuracy for MI + MNS vs. MNS task, but also to optimize the proposed method by (i) selecting the best frequency band, (ii) selecting the best feature extraction and classifier method and its parameters, and (iii) reducing the number of EEG electrodes to use.

II. Materials and Methods

A. Participants

Twenty-two right-handed healthy volunteers (12 females; 19 to 57 years old; 28.56 $\pm$ 13.3 years old) were recruited for this study. All subjects satisfied the inclusion criteria (right-handed, between 18-60 years old, without medical history which could have influenced the task). This experiment followed the statements of the WMA Declaration of Helsinki on ethical principles for medical research involving human subjects [24]. In addition, participants signed an informed consent which was approved by the ethical committee of Inria (COERLE, approval number: 2016-011/01) as it satisfied the ethical rules and principles of the institute.

B. Experimental task

During EEG recordings the subject was lying on a comfortable chair with eyes closed. Their legs rest on a footrest and the right forearm rests on a cushion to prevent movement. The subject mentally presses and releases a remote button. In previous works [10], [25], we showed that even such a minimalist movement could be sufficient to modulate the cerebral motor activity, especially in terms of ERD and ERS modulations in the EEG signal and be detected. We also showed that more complex intentions of movements such as combined motor imageries can also be detected [26]. The operator monitors the EEG signals during the experiment.

The aim of this research is to detect the occurrence of motor patterns under two different conditions: median nerve stimulation during motor imagery (MI + MNS) and median nerve stimulation during rest (MNS). Participants performed each condition randomly during two runs of 26 trials per condition. So, in total, the database consists of 52 recordings of each class for every subject.

Condition 1: motor imagery:

For the MI condition (C1), subjects had to imagine an isometric grasp between the thumb and the index finger on a pointer button, i.e., they had to try to feel a maximum of sensations caused by the real movement, but without any movement. A low-frequency beep indicated when the subject had to start the

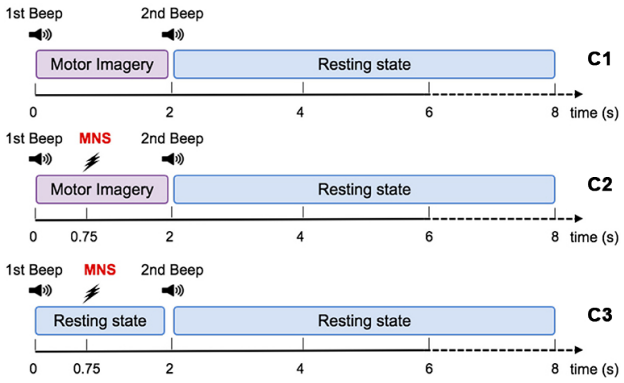


Fig. 1: Representation scheme for one trial. Timing schemes of a trial for C1, C2, and C3. For all motor tasks, one low-frequency beep indicates when to start the task. For the MNS+MI condition, the MNS occurs at 750 milliseconds after the first beep. The end of the MI is announced by a high-frequency beep and followed by a rest period of 6 seconds.

motor imagery, the grasping MI was maintained during 2 s, then a second beep indicated the end of the imagined task (Fig. 1).

Condition 2: motor imagery with a median nerve stimulation: During the MI + MNS condition (C2), subjects had to perform the same motor imagination as C1 while their median nerve was stimulated at 750 milliseconds after the start of the motor imaginary task (Fig. 1). A low-frequency beep was indicated when the subject had to start the motor imagery and a high-frequency beep indicated when to stop it. We chose the 750 milliseconds of delay according to the reaction time of the average person, in order to stimulate during the ERD corresponding to the imagination started. The duration of stimulation was 100 ms.

Condition 3: median nerve stimulation only: The MNS only condition (C3) consisted of a series of stimulation of the median nerve during rest (Fig. 1). We placed two stimulating electrodes on the wrist according to the standards [15], [27] and recorded the EEG data from subjects.

C. Electrophysiological data

EEG signals were acquired using the *OpenViBE* platform with a *BioSemi Active Two* 128-channel EEG system, arranged in the *BioSemi*'s ABC system covering the entire scalp at 2048Hz. Among all registered sites, some of the electrodes were localized around the primary motor cortex, motor cortex, somatosensory cortex, and occipital cortex, which allowed to observe the physiological changes due to kinesthetic motor imagery and median nerve stimulation [12]. An external electromyogram electrode was added in order to verify that there was no movement during the MI task.

D. EEG channel layouts

Initially 128 EEG channels were registered with *BioSemi*'s ABC layout. The recordings with a reduced number of channels were constructed to investigate the possible influence of using a smaller number of electrodes on the motor imagery detection performance. We also investigated additional channels from the frontal lobe as an interesting area for our experiment because it is considered as the emotional controller center and also responsible for motor functions, motor speech functions and impulse control [28]. Electrode layouts tested in the current work represented in Fig. 2 and in Table I.

E. Data preprocessing

After selecting a certain channel set, EEG signals were re-referenced to common average reference, then EEG signals were resampled at 128 Hz.

Then each signal was filtered in a specific frequency band using a 4th-order Butterworth band-pass filter [29] and divided into trials of 3 s duration starting at 250 ms after the first beep (Fig. 1). We investigated filtering in few frequency bands to see is it efficient enough to add theta (4-8 Hz) or low gamma (30-38 Hz) range to standard mu+beta (8-30 Hz) frequency band especially for MNS vs. MI+MNS discriminating task because it has not been studied before. Thereby we used the following bands: theta-beta (4-30 Hz), standard mu-beta (8-30 Hz), mu-gamma (8-38 Hz), and theta-gamma (4-38 Hz) [30].

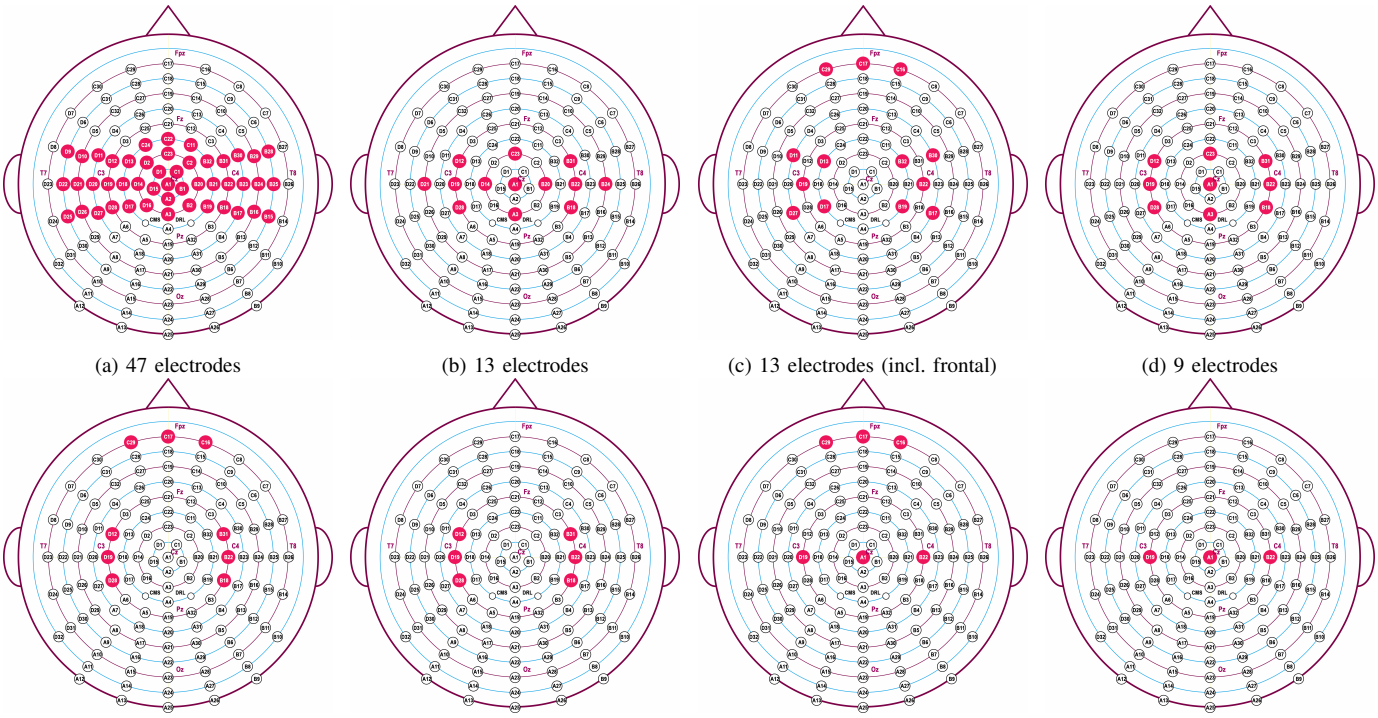
F. Feature extraction and classification

In this study, we selected three state-of-the-art deep neural network architectures to compare their results to the best-known approaches for classification in the MI-based BCI domain. For feature extraction we used two efficient methods: CSP filters [16] and Riemannian geometry [17], [18]. For classification we used first a linear discriminant analysis (LDA). We also selected approaches based on Riemannian geometry: MDRM algorithm [17], [18], tangent space projection of covariance matrices with following logistic regression classifier application on projected data (TS+LR) as described in [10], [17].

Our motivation to use convolutional neural networks is to not predefine features and expect better accuracy. We used i) a deep convolutional network (DeepConvNet) [21]) as reference for deep networks, ii) a shallow convolutional network (ShallowConvNet) [21] because they are designed specifically for oscillatory signal classification, and iii) a generic architecture (EEGNet) proposed in [22] for EEG data that already demonstrated good results on several BCI paradigms.

We did not use any feature extraction before deep learning algorithms; feeding filtered EEG signals as an input to them.

DeepConvNet: DeepConvNet has four convolution-max-pooling blocks, with a special first block designed to handle the EEG input, followed by three standard convolution-max-pooling blocks and a dense softmax classification layer. The first convolutional block was split into two convolutional layers in order to better handle the large number of input channels — one input channel per electrode [21]. Full details about the network architecture used in this paper can be found in Table II.



(a) 47 electrodes (b) 13 electrodes (c) 13 electrodes (incl. frontal) (d) 9 electrodes

(e) 9 electrodes (incl. frontal) (f) 6 electrodes (g) 6 electrodes (incl. frontal) (h) 3 electrodes

Fig. 2: Layouts with electrode positions tested in comparison to the 128 electrodes montage.

ShallowConvNet: The architecture designed for the Shallow Convolutional Network has been inspired by the Filter Bank Common Spatial Pattern (FBCSP) algorithm. The transformations performed by the ShallowConvNet are similar to the transformations of FBCSP. Concretely, the first two layers of the ShallowConvNet perform a temporal and a spatial convolution, as in the DeepConvNet. These steps are analogous to the bandpass filter and the CSP spatial filter steps in FBCSP. In contrast to the DeepConvNet, the temporal convolution of the ShallowConvNet had a larger kernel size, allowing a larger range of transformations in this layer. After the two convolutions of the ShallowConvNet, a squaring nonlinearity x^2 activation function, a mean pooling layer and a logarithmic activation function $log(x)$ followed; together these steps are analogous to the trial log-variance computation in FBCSP [21]. Full details about the network architecture can be found in Table III.

EEGNet: EEGNet architecture is presented on Fig. 3. The first layer of the EEGNet is a temporal convolution to learn frequency filters. Then follows a depthwise convolution layer, connected to each feature map individually to learn frequency-specific spatial filters. The separable convolution is the next convolutional layer; it is a combination of a depthwise convolution, which learns a temporal summary for each feature map individually, followed by a pointwise convolution, which learns how to optimally mix the feature maps together [22].

We tested different sets of parameters (not detailed in the paper) and select 8 models derived from the original EEGNet which allow us to obtain the best result. We kept 8 temporal filters $F_1 = 8$ and 16 pointwise filters $F_2 = 16$ for all models presented in this paper. The parameters which have been changing were lengths of temporal kernel size for the first convolutional layer K and multiplier depth D (which determines the number of spatial filters) of the depthwise convolution layer. Thus, we set multiplier depth to 2 or 4.

The first layer of filters of the original EEGNet [22] extracts mainly a low frequency information since it has large kernel size $(1, 32)$, and ignores mid- and high frequency information [31]. To remedy this problem, we reduced the first layer kernel size (as in [31]). So for this study the kernel size was set to $(1, 32)$ (as recommended in [22] for sensory motor data), $(1, 16)$, $(1, 8)$ and $(1, 4)$. These new architectures retain much more information in the first layer features concerning the mid and high-frequency bands. More importantly, they also improve the classification performance as shown in Section III-B. We use the notation EEGNet-D.K to denote the number of spatial filters to learn and kernel length for the first convolutional layer; i.e., EEGNet-4.8 denotes learning 4 spatial filters per temporal filter and kernel length $(1, 8)$. Full details about the network architecture can be found in Table IV.

Training procedure: Deep neural networks were implemented with the TensorFlow framework and Keras [32] and trained using the Grid'5000 testbed [33] to speed up the computations. The training of the networks is performed with the following configurations.

- ADAM [34] is used as the optimization method. The parameters are set to default values as proposed in [34].
- Categorical cross-entropy is used as the optimization criterion.
- Batch size is selected as 16.
- We run 300 training iterations (epochs) and perform validation stopping, saving the model weights which

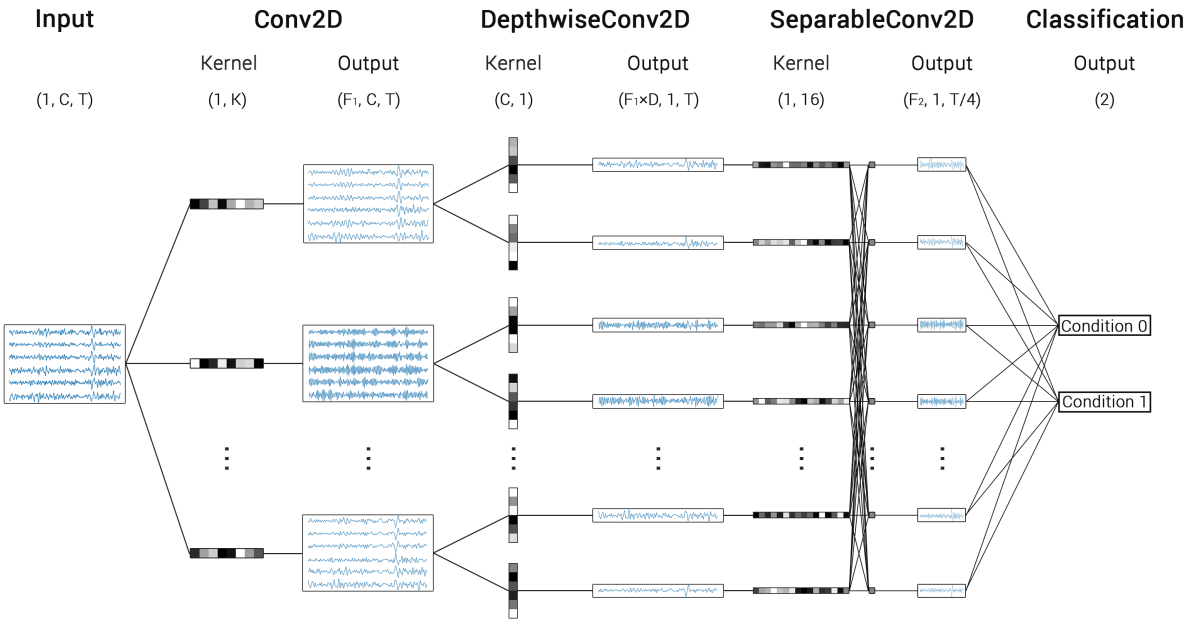


Fig. 3: EEGNet architecture. Where the number of channels $C = 6$, the number of time samples T, kernel length for the first convolutional layer $K = 8$, multiplier depth $D = 2$, the number of temporal filters F_1 and the number of pointwise filters F_2. Here condition 1 corresponds to the MI or MI+MNS class and condition 0 corresponds to the rest or respectively to MNS class depending on the task.

produced the lowest validation set loss.

- We used a batch normalization [35] layer before the activation layer and a dropout layer after the activation with a probability of 50%. Our investigations showed that 50% dropout probability leads to better (approx. 2%) performance comparing to 25% (as proposed in [22]) and helps to prevent overfitting on datasets with a small number of samples.

As a threshold accepted in the present study, 80% of accuracy and 20% of false-positive rate (FPR) was used. It is recommended to use the FPR metric which is very important in the task of detecting intraoperative awareness as it is responsible for false alarms which are more than undesirable in this task.

G. Topographies

To analyze the differences between MI + MNS and MNS conditions, we performed an event-related spectral perturbation (ERSP) analysis for two different frequency bands (8-30 Hz vs. 4-38 Hz) using EEGLAB (Fig. 4a). It allowed us to understand how MI + MNS and MNS conditions can be discriminated and which time parameters we can choose to guide the classification (Fig. 4b). A surrogate permutation test ($p < 0.01$; 4,000 permutations) from the EEGLAB toolbox was used to validate differences in terms of localization of these ERSPs. In addition to this analysis, we applied a false discovery rate (FDR) correction test in order to clarify how the FDR was controlled for multiple comparisons.

H. Cross-validation and statistical analysis

We report subject-specific classification results corresponding to ten-fold blockwise cross-validation, where eight of ten blocks are chosen as the training set, one block as the validation set, and the remaining block as the test set. We performed a Wilcoxon signed-rank test (two-sided) to find a significant difference in accuracy obtained by different models. We report different significance levels ($p - value < 0.05, p - value < 0.01, p - value < 0.001$) in the plots.

III. Results

In this result section, we present how deep neural nets architectures can improve the classification accuracy for two MI discriminating paradigms: the standard (MI vs. Rest) BCI task and the same task when stimulation of the median nerve is added (MI+MNS vs. MNS). We addressed the following aspects of the experiments:

- using extended frequency ranges with theta and low gamma bands (Section III-A);
- using various deep learning architectures and standard classifiers (Section III-C);
- comparing the two discriminating tasks: MNS vs. MNS+MI and MI vs. Rest (Section III-D).
- influence of different EEG montages (Section III-E);

A. Frequency band

In the first set of experiments, we investigate the influence of the EEG frequency band on the classification accuracy. The recommended frequency band for detecting motor imagery is 8-30 Hz [36], but sometimes it could be useful to extend the analysis to wider frequency bands such as 4-38 Hz [37]. We investigated separately and together the impact of adding the theta 4-8 Hz and the low gamma 30-38 Hz frequency ranges on the classification performance. We did this study for all models, i.e., CSP+LDA, MDRM, TS+LR, EEGNet, DeepConvNet, and ShallowConvNet. Similar behavior was obtained for

all models, so we only present the most representative part of it (see subsection III-C) on Fig. 5 showing classification accuracies for EEGNet-4.8 model trained on data filtered with different bands for MNS vs. MNS+MI and MI vs. Rest tasks.

1) MNS+MI vs. MNS: Fig. 5a shows that including separately and both together the theta and the low gamma bands significantly ($p-value < 0.01$) improves the accuracy. Indeed, discrimination power using 4-30 Hz and 8-38 Hz bands exceeds the one of the 8-30 Hz band. Moreover, both together they improve even more accuracy. The positive effect on the accuracy with the low gamma band is stronger ($p - value < 0.001$) for 128 electrodes.

Wilcoxon signed-rank test shows significant ($p - value < 0.01$ or higher) difference between EEGNet performance on data filtered in 8-30 Hz and 4-38 Hz frequency bands for MNS vs. MI+MNS discriminating task. This is the case for 128, 47, 13, 9, and 6-electrode setups. At the same time, for 47, 13, and 6-electrode setups the difference was even more pronounced ($p - value < 0.001$). The advantage of using a 4-38 Hz band can achieve accuracy improvement up to 4% for MNS vs. MI+MNS task.

2) MI vs. Rest: Although for the MI vs. Rest task the addition of the theta band influences classification accuracy, but no significant difference has been noticed based on Wilcoxon signed-rank test (Fig. 5b).

3) Recommendations: When using MNS it is better to use the wider frequency band 4-38 Hz. At the same time, there are no reasons to change the standard 8-30 Hz band for the Rest vs. MI task.

B. EEGNet architectures

EEGNet is a very promising architecture [22] for spatio-temporal analysis of EEG signals. To better investigate this architecture, many parameters of EEGNet may be adjusted. In this work, we focus on the temporal kernel size K for the first convolutional layer and the multiplier depth D of the depthwise convolution layer which were not investigated deeply earlier. The accuracies obtained using different architectures of EEGNet for MNS vs. MI+MNS task are shown in Fig. 6a. We can see that decreasing the kernel size of the first convolutional layer and increasing the multiplier depth outperforms the standard set of parameters of EEGNet-2.32 [22] for a kernel size of 2 and a multiplier depth of 32. We selected EEGNet-4.8 for future investigations based on its higher classification performance for almost all electrode setups.

More precisely comparing EEGNet-4.8 proposed in this work with standard EEGNet-2.32 architecture for sensory-motor data [22] we can see that it works especially better for larger setups of electrodes for MNS vs. MI+MNS task (Figure 6a). Wilcoxon signed-rank test showed a highly significant ($p - value < 0.01$) difference between performances of EEGNet-4.8 and EEGNet-2.32 for all setups of electrodes except setup with 3 electrodes over the motor cortex. For MI vs. Rest task (Fig. 6b) the difference is not statistically significant, but the classification accuracy of EEGNet-4.8 is better for most electrode setups.

C. Classification results

Classification accuracies obtained by EEGNet and other standard and deep learning models are presented in Fig. 7a for MI + MNS vs. MNS task and in Fig. 7b for MI vs. Rest task.

1) MI + MNS vs. MNS task: For this task EEGNet-4.8 significantly ($p - value < 0.01$) outperforms all standard and deep learning models for all electrode setups, except 3-electrode setup where difference to second best model (EEGNet-2.32) is not significant (Fig. 7a). The second best model for all electrodes setups is also an EEGNet architecture (EEGNet-2.32).

Also from Fig. 7a we can see those performances of DeepConvNet and ShallowConvNet become better when the electrode number decreases from 128 to 47, whereas it becomes worse for all other models. All deep learning models outperform standard models starting from setups with 47 electrodes and less.

2) MI vs. Rest task: TS+LR outperforms other models for 128 and 47 electrodes for MI vs. Rest task (Fig. 7b) but then performance decreases a lot when decreasing the number of electrodes. At the same time, deep learning models are more stable to decreasing of electrodes number. And we even observe that a decreasing number of electrodes leads to increasing classification performance for all deep learning models with a peak for the 9 electrode setup. Statistical tests show that EEGNet-4.8 significantly ($p - value < 0.05$) outperforms TS+LR (and therefore other standard methods) for 9 and 6 electrode setups.

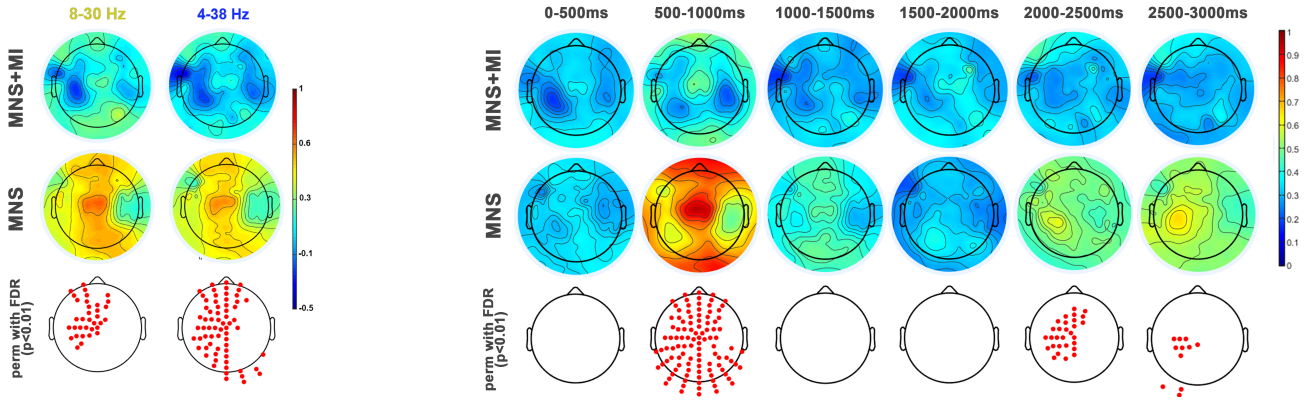
D. Comparing paradigms

The difference between results of MI + MNS vs. MNS and MI vs. Rest paradigms was demonstrated earlier in [10], [23]. But previous investigations were done only for the 128 electrode setup. In this work, we also present it for a smaller number of electrodes. Boxplots with classification accuracy distribution for EEGNet-4.8 model for setups with 13, 9, 6 electrodes located over motor cortex are shown in Fig. 8. Statistical tests indicate highly significant ($p < 0.001$) difference between EEGNet-4.8 accuracies of both tasks for 13, 9 and 6 electrodes.

E. Frontal electrodes investigation

To explore the opportunities of including the EEG activity from different functional areas of the brain, we investigated setups with electrodes not only from motor cortex but also adding electrodes from frontal lobe (Figures 2c, 2e, 2g) because frontal electrodes are used in many monitoring systems to determine the depth of anesthesia [6].

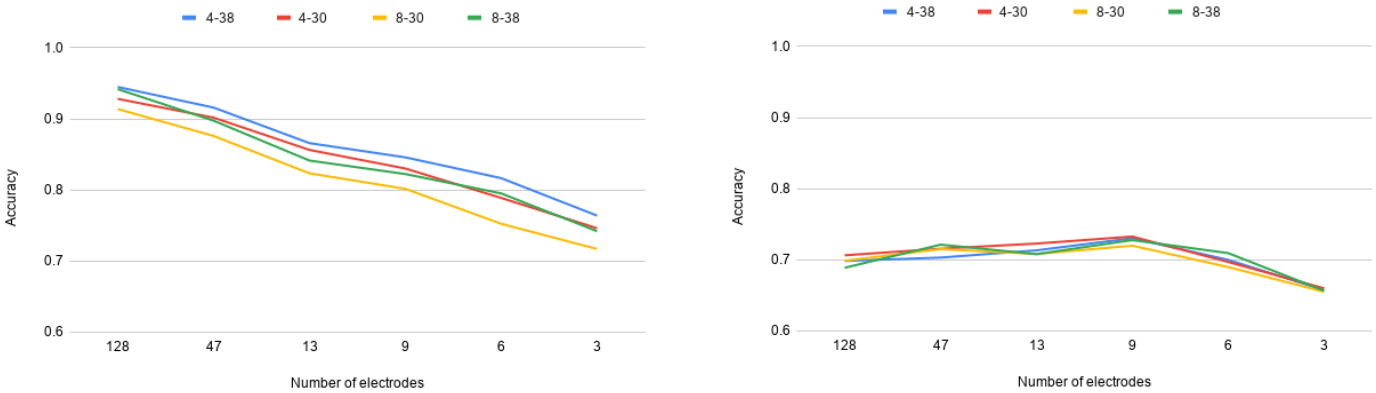
Fig. 9 shows classification performance of models for MI + MNS vs MNS discriminating task on the data from 13, 9, and 6 electrodes over the motor cortex and frontal lobe (distinguished by prefix $+fr$ in the chart), and motor cortex only. Results show that the addition of frontal lobe electrodes increases accuracy almost for all models and all numbers of electrodes. Wilcoxon signed-rank tests confirm significant



(a) 8-30 Hz and 4-38 Hz band comparison.

(b) Dynamics of ERD/ERS% over the 3s period of analysis in the 4-38 Hz band.

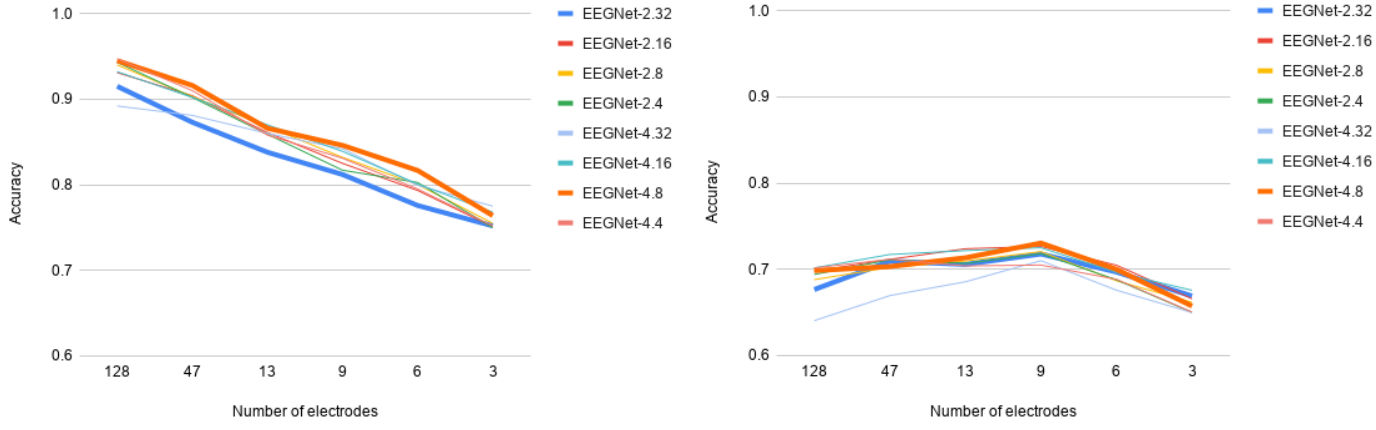
Fig. 4: Topographic map of ERD/ERS% (grand average, n = 22) during two conditions: MI + MNS (C2) and MNS only (C3). A red color corresponds to a strong ERS and a blue one to a strong ERD. Red electrodes indicate a significant difference between the two conditions ($p < 0.01$).



(a) EEGNet-4.8 for MI + MNS (C2) vs. MNS (C3) task

(b) EEGNet-4.8 for MI (C1) vs. Rest task

Fig. 5: Classification accuracy for EEGNet-4.8 averaged over all folds ($k = 10$) and all subjects ($n = 22$) for different electrode layouts of MI + MNS (C2) vs. MNS (C3) (5a) and MI (C1) vs. Rest (5b) discriminating task. Input data filtered in 4-38 Hz, 4-30 Hz, 8-30 Hz and 8-38 Hz frequency ranges.

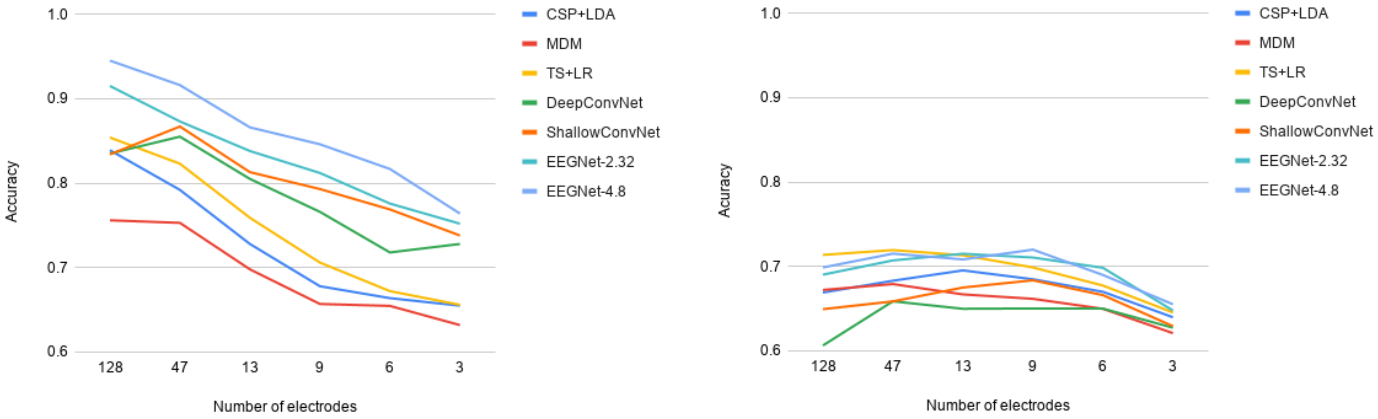


(a) MI + MNS (C2) vs. MNS (C3) task

(b) MI (C1) vs. Rest task

Fig. 6: Classification accuracy for different variants of EEGNet, averaged over all folds ($k = 10$) and all subjects ($n = 22$) for different electrode layouts of MI + MNS (C2) vs. MNS (C3) (6a) and MI (C1) vs. Rest (6b) discriminating task. Input data filtered in 4-38 Hz frequency band for 6a and in 8-30 Hz frequency band for 6b.

differences between setups with and without frontal lobe electrodes for many models. For example, the difference is significant ($p - value < 0.05$) for TS+LR and EEGNet-2.32 for setups with 9 and 6 electrodes. Moreover, it is highly significant for the TS+LR model using 6 electrodes. At the same time, statistical tests show that there is no



(a) MI + MNS (C2) vs. MNS (C3) task

(b) MI (C1) vs. Rest task

Fig. 7: Classification accuracy for different electrode setups, averaged over all folds ($k = 10$) and all subjects ($n = 22$) for MI + MNS (C2) vs. MNS (C3) (7a) and MI (C1) vs. Rest (7b) conditions. Input data filtered in 4-38 Hz frequency band for 7a and in 8-30 Hz frequency band for 7b.

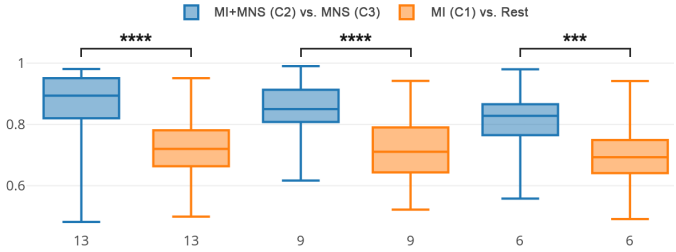


Fig. 8: Boxplots showing the distribution of EEGNet-4.8 classification accuracy for different setups of n electrodes over motor cortex and frontal lobe, averaged over all folds ($k = 10$) and all subjects ($n = 22$) for both discriminating tasks. Significance levels of Wilcoxon signed-rank test comparisons $^{***}p < 0.001$, $^{****}p < 0.0001$.

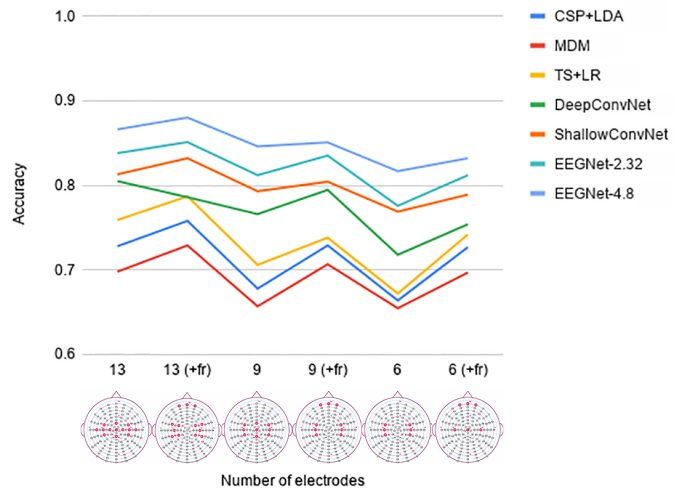


Fig. 9: Models classification accuracy for electrode setups over the motor cortex and over the motor cortex and frontal lobe, averaged over all folds ($k = 10$) and all subjects ($n = 22$) for MI + MNS (C2) vs. MNS condition (C3).

significant difference between setups with and without frontal lobe electrodes for EEGNet-4.8.

IV. Discussion

In this work we addressed the problem of MI detection from raw EEG. In our considered case, the subject's median nerve is being stimulated while MI is being detected, which has a potential application in anesthesia awareness detection. We demonstrated that for such a setup the application of modified EEGNet deep neural network to the EEG in extended frequency band results in high classification accuracy (83.2%) and low false positive rate (19.0%) for a setup with only six electrodes over motor cortex and frontal lobe.

A. Different deep learning architectures

We tested four different deep learning models: DeepConvNet, ShallowConvNet, EEGNet-2.32, and our recommended modification of EEGNet (EEGNet-4.8) where we reduced the first layer kernel size from 32 to 8 to better handle mid- and high-frequency features. ShallowConvNet and EEGNets showed high performances. Statistical tests show that EEGNet-4.8, proposed in this paper, significantly outperformed all standard and deep learning approaches especially for MNS + MI vs. MNS task in terms of accuracy.

The higher performance of EEGNet-4.8 compare to EEGNet-2.32 for MNS + MI vs. MNS task can be explained by the presence of useful high-frequency components in the EEG signals for this task unlike MI vs. Rest task. These components are better kept using a shorter kernel size. It is also shown in Fig. 5, where we can see better performances for 8-38 Hz than for standard 8-30 Hz frequency band. Usually, there is intuition to use smaller kernels for detecting high-frequency features and larger kernels for low-frequency features [31]. That's why reducing the first layer kernel size from 32 to 8 is useful for MNS + MI vs. MNS task.

Also, as we use the same kernel for a different set of samples in a signal, the same weights are shared across these sample sets. This means that a smaller kernel size leads to using fewer weights to back-propagate on. So the benefit of using a smaller kernel is a reduction in computational costs. We also changed multiplier depth from 2 to 4 in the second layer to increase

the number of spatial filters, which allows us to store more information in this layer.

B. Frequency bands

We found that using a 4-38 Hz frequency band is helpful for better distinguishing motor patterns resulting in MNS + MI and MNS. Previously, the 8-30 Hz frequency range was mainly used in BCI [36], but adding the theta and low gamma frequency ranges resulted in an increase of accuracy for each electrode layout studied in our experiments.

On one hand, the effect on hand motor function in the theta 4.8–7.9 Hz band of somatosensory electrical stimulation has already been shown [38]. They observed a significant decrease in mean motor ipsilesional resting-state power.

On the other hand, the are many confirmations of motor activities in the gamma band [39], [40].

C. Frontal electrodes

When investigating the effect of adding channels from the frontal cortex as additional input data to the models, we noticed that for many models the accuracy increases comparing to the case when only the channels from the motor cortex were used. This was the case for the less accurate models (e.g. TS+LR and EEGNet-2.32), whereas there were no significant improvements for the best-obtained model (EEGNet-4.8). The evident increase in classification accuracy between MI + MNS vs. MNS after adding frontal electrodes might have different causes, leading to models perceiving the presence of MI conditioned to nerve stimulation. Possible reasons may be rooted in participation of the prefrontal cortex in response implementation, providing a functional substrate for the transformation of perception into action [41], or regions of the frontal cortex being a part of the cortical networks involved in the complex experience of pain [42].

Although current results show that the proposed EEGNet-4.8 model with the highest accuracy demonstrates that the inclusion of the frontal electrodes does not lead to even more statistically significant improvements in terms of accuracy, we suggest to use frontal electrodes for discriminating MI + MNS and MNS since adding them statistically significantly improves the performance for most of the classifiers (Fig. 9). Therefore, we believe that the potential of looking at the frontal area is high and this direction will be addressed in our future research. Information from the frontal cortex is useful and the ways of integrating it into classification models in order to obtain a more significant increase in accuracy should be explored further.

In parallel with testing the performance of many models, we addressed the problem of minimization of the electrode number in the recording setup, to reduce the preparation and calibration time in clinical settings. It is especially important for the task of detecting intraoperative awareness, where preparation time before surgery is a precious resource. Starting from 128 electrodes layout, we tested the dependency of the performance on the electrode number for 48, 13, 9, 6, and 3 electrodes covering mainly the motor cortex area. We found that it is possible to reduce the number of electrodes to 13 or even 6 electrodes without substantial loss of classification performance. The best model (EEGNet-4.8) achieved 88% accuracy for 13 electrodes setup (83.2% for 6 electrodes setup) for MNS + MI vs. MNS task. Since such a task with a few electrodes was not addressed in the literature previously, this is the highest accuracy to the best of the authors' knowledge. These values are higher than the available values for the model which distinguishes MI vs. Rest, which is 73.8% for 13 electrodes setup (68.9% for 6 electrodes setup). Deep learning models showed higher stability when decreasing the number of electrodes for MNS + MI vs. MNS task. At the same time, the best classification performance for MI vs. Rest task has been achieved on the setup with 9 electrodes.

D. Intraoperative awareness

The methodology presented in this article highlights the benefits of adding a median nerve stimulation (with an adapted 8-38 Hz frequency band and frontal electrodes) to increase the detection of a motor intention in particular using adapted deep learning architectures. This approach is also very interesting in the context of AAGA since a frequent stimulation can play the role of a trigger to precise when analyzing a patient's EEG. Nevertheless, under real conditions of a surgery, EEG patterns of motor intention under anesthetic agent, such as propofol, are not available since no sample can be previously labeled. Then, further investigations will be done to tackle this problem using EEG patterns of motor intention collected before the induction of propofol and using transfer learning methods [43] or one-class approaches. Although most anesthetics reduce the global metabolic activity of the central nervous system by decreasing neuronal excitability [44], the effect of anesthetics on the sensorimotor cortex appears to be relatively well preserved [9], [45]. In fact, regions of the cortex would still be receptive to information, but their ability to communicate with other regions seems to be damaged during general anesthesia [46], suggesting that our method can be used with patients under clinical conditions [9].

V. Conclusion

This study confirms that deep learning networks can outperform state-of-the-art classification methods to detect motor imagery especially since they only use filtered data and not well-known extracted features. It also confirms that adding median nerve stimulation (MNS) facilitates this detection. Moreover, we have shown that signals recorded by frontal electrodes and a larger frequency band 4-38 Hz contain relevant information when adding median nerve stimulation to discriminating right hand motor imagery from MNS only. These results confirm the effect of MNS on the entire motor system. Thus, we demonstrated that for detection of motor imagery it is advisable to use EEG recorded from as few as 6 electrodes in 4-38 Hz frequency range processed by EEGNet-4.8 deep neural network while the subject is being stimulated via the median nerve, which can give the 83.2% of classification accuracy and 19.0% of false-positive rate. Combination of the novel approach of MNS with a larger number of electrodes results in higher accuracy (94.5% for 128 electrodes and 88.0% for 13

electrodes) whereas acceptable FPR (6.1% for 128 electrodes and 12.9% for 13 electrodes). These results set the background for clinical application with anesthetics of the proposed setup for accessing the intraoperative awareness.

REFERENCES

[1] T. Weiser, *et al.*, "Size and distribution of the global volume of surgery in 2012," *Bull World Health Organ*, vol. 94, no. 3, pp. 201–209, 2016.

[2] G. Mashour and M. Avidan, "Intraoperative awareness: controversies and non-controversies," *British Journal of Anaesthesia*, vol. 115, pp. 20–26, 2015.

[3] D. Almeida, "Awake and unable to move: what can perioperative practitioners do to avoid accidental awareness under general anesthesia?" *J Perioper Pract*, vol. 25, no. 12, pp. 257–261, 2015.

[4] R. Blacher, "On awakening paralyzed during surgery. a syndrome of traumatic neurosis," *Br J Anaesth*, no. 234, pp. 67–8, 1975.

[5] C. Kent and K. Domino, "Depth of anesthesia," *Curr opin Anaesthesiol*, vol. 22, no. 6, pp. 782–787, 2009.

[6] P. L. Purdon, *et al.*, "Clinical Electroencephalography for Anesthesiologists Part I: Background and Basic Signatures," *Anesthesiology*, vol. 123, no. 4, pp. 937–960, Oct. 2015.

[7] Y. Punjasawadwong, *et al.*, "Bispectral index for improving anaesthetic delivery and postoperative recovery," *Cochrane Database of Systematic Reviews*, no. 6, 2014.

[8] S. Rimbert, *et al.*, "Motana: study protocol to investigate motor cerebral activity during a propofol sedation," *Trial*, vol. 20, no. 534, p. 9, Sept. 2019.

[9] S. Rimbert, *et al.*, "Innovative brain-computer interface based on motor cortex activity to detect accidental awareness during general anesthesia," in *IEEE SMC 2019 - IEEE International Conference on Systems, Man, and Cybernetics*, Bari, Italy, Oct. 2019.

[10] S. Rimbert, *et al.*, "Median Nerve Stimulation Based BCI: A New Approach to Detect Intraoperative Awareness During General Anesthesia," *Frontiers in Neuroscience*, vol. 13, p. 622, June 2019.

[11] M. M. Ghoneim, *et al.*, "Awareness during anesthesia: risk factors, causes and sequelae: a review of reported cases in the literature," *Anesthesia and Analgesia*, vol. 108, no. 2, pp. 527–535, 2009.

[12] S. Salenius, *et al.*, "Modulation of human cortical rolandic rhythms during natural sensorimotor tasks," *NeuroImage*, vol. 5, no. 3, pp. 221–228, 1997.

[13] C. Neuper and G. Pfurtscheller, "Event-related dynamics of cortical thythms: frequency-specific features and functional correlates," *Int J Psychophysiology*, vol. 43, pp. 41–58, 2001.

[14] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clin Neurophysiol*, vol. 110, no. 11, pp. 1842–57, Nov. 1999.

[15] A. Schnitzler, *et al.*, "Involvement of primary motor cortex in motor imagery: a neuromagnetic study." *Neuroimage*, vol. 6, no. 3, pp. 201–208, Oct. 1997.

[16] Y. Fang, *et al.*, "Extending csp to detect motor imagery in a four-class bci," *Journal of Information & Computational Science*, vol. 9, no. 1, 2012.

[17] A. Barachant, *et al.*, "Riemannian geometry applied to bci classification," in *International Conference on Latent Variable Analysis and Signal Separation*. Springer, 2010, pp. 629–636.

[18] M. Congedo, *et al.*, "A New Generation of Brain-Computer Interface Based on Riemannian Geometry," *arXiv:1310.8115 [cs, math]*, Oct. 2013, arXiv: 1310.8115.

[19] Y. Roy, *et al.*, "Deep learning-based electroencephalography analysis: a systematic review," *Journal of Neural Engineering*, vol. 16, no. 5, p. 051001, Aug. 2019.

[20] A. Craik, *et al.*, "Deep learning for electroencephalogram (EEG) classification tasks: a review," *Journal of Neural Engineering*, vol. 16, no. 3, p. 031001, June 2019.

[21] R. T. Schirrmeister, *et al.*, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, Nov. 2017, arXiv: 1703.05051.

[22] V. J. Lawhern, *et al.*, "EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces," *Journal of Neural Engineering*, vol. 15, no. 5, p. 056013, July 2018.

[23] O. Avilov, *et al.*, "Deep learning techniques to improve intraoperative awareness detection from electroencephalographic signals," *Conference proceedings: 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) in conjunction with the 43rd Annual Conference of the Canadian Medical and Biological Engineering Society. July 20-24, 2020, Montréal, Canada (forthcoming)*, 2020.

[24] W. M. Association, "World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects." *Journal of Postgraduate Medicine*, vol. 48, no. 3, p. 206, Jan. 2002, publisher: Medknow Publications.

[25] S. Rimbert, *et al.*, "Profiling BCI users based on contralateral activity to improve kinesthetic motor imagery detection," *IEEE Neural Engineering (NER) 2017*, 2017.

[26] C. Lindig-León, *et al.*, "Multiclass Classification Based on Combined Motor Imageries," *Frontiers in Neuroscience*, vol. 14, Nov. 2020.

[27] D. Kumbhare, *et al.*, *Buschbacher's Manual of Nerve Conduction Studies, 3rd Edition*. Demos Medical Publishing New york, 2016, ch. Mediane nerve to the abductor pellicis brevis, p. 10.

[28] A. Collins and E. Koechlin, "Reasoning, Learning, and Creativity: Frontal Lobe Function and Human Decision-Making," *PLoS Biology*, vol. 10, no. 3, Mar. 2012.

[29] S. Butterworth, "On the theory of filter amplifiers," *In Wireless Engineer (also called Experimental Wireless and the Wireless Engineer)*, vol. 7, p. 536–541, 1930.

[30] A. Baghdadi, *et al.*, *A Survey of Methods and Performances for EEG-Based Emotion Recognition*, ser. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, 2017, pp. 164–174.

[31] M. D. Zeiler and R. Fergus, "Visualizing and Understanding Convolutional Networks," in *Computer Vision – ECCV 2014*, D. Fleet, *et al.*, Eds. Cham: Springer International Publishing, 2014, vol. 8689, pp. 818–833, series Title: Lecture Notes in Computer Science.

[32] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[33] D. Balouek, *et al.*, "Adding virtualization capabilities to the Grid'5000 testbed," in *Cloud Computing and Services Science*, ser. Communications in Computer and Information Science, I. I. Ivanov, *et al.*, Eds. Springer International Publishing, 2013, vol. 367, pp. 3–20.

[34] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv:1412.6980 [cs]*, Dec. 2014, arXiv: 1412.6980.

[35] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *arXiv:1502.03167 [cs]*, Mar. 2015, arXiv: 1502.03167.

[36] F. Lotte, *et al.*, "Comparison of designs towards a subject-independent brain-computer interface based on motor imagery," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Minneapolis, MN: IEEE, Sept. 2009, pp. 4543–4546.

[37] K.-W. Ha and J.-W. Jeong, "Motor Imagery EEG Classification Using Capsule Networks," *Sensors (Basel, Switzerland)*, vol. 19, no. 13, June 2019.

[38] A. P. Tu-Chan, *et al.*, "Effects of somatosensory electrical stimulation on motor function and cortical oscillations," *Journal of NeuroEngineering and Rehabilitation*, vol. 14, no. 1, p. 113, 2017.

[39] G. Buzsáki and X.-J. Wang, "Mechanisms of gamma oscillations," *Annual Review of Neuroscience*, vol. 35, pp. 203–225, 2012.

[40] M. Ahn, *et al.*, "Gamma band activity associated with BCI performance: simultaneous MEG/EEG study," *Frontiers in Human Neuroscience*, vol. 7, 2013.

[41] M. Haller, *et al.*, "Persistent neuronal activity in human prefrontal cortex links perception and action," *Nature Human Behaviour*, vol. 2, no. 1, pp. 80–91, Jan. 2018, number: 1 Publisher: Nature Publishing Group.

[42] M. C. Bushnell, *et al.*, "Pain perception: Is there a role for primary somatosensory cortex?" *Proceedings of the National Academy of Sciences*, vol. 96, no. 14, pp. 7705–7709, July 1999, publisher: National Academy of Sciences Section: Colloquium Paper.

[43] L. Bougrain, *et al.*, "Guidelines to use Transfer Learning for Motor Imagery Detection: an experimental study," in *10th International IEEE/EMBS Conference on Neural Engineering (NER).*, virtually, United States, May 2021.

[44] G. A. Mashour and R. Lydic, "Neuroscientific Foundations of Anesthesiology," *European Journal of Anaesthesiology (EJA)*, vol. 29, no. 6, p. 301, June 2012.

[45] Y. Blokland, *et al.*, "Decoding motor responses from the eeg during altered states of consciousness induced by propofol," *Journal of Neural Engineering*, vol. 13, no. 2, p. 026014, 2016.

[46] S. J. Peltier, *et al.*, "Functional connectivity changes with concentration of sevoflurane anesthesia," *NeuroReport*, vol. 16, no. 3, pp. 285–288, Feb. 2005.

APPENDIX

TABLE I: Investigated electrode layouts

Layout	Electrodes	Figure
128 channels	all available electrodes	
47 channels	A1, A2, A3, B1, B2, B15, B16, B17, B18, B19, B20, B21, B22, B23, B24, B25, B28, B29, B30, B31, B32, C1, C2, C23, C24, C22, C11, D1, D2, D9, D10, D11, D12, D13, D14, D15, D16, D17, D18, D19, D20, D21, D22, D25, D26, D27, D28	Fig. 2a
13 channels over the motor cortex	D12, C23, B31, D21, D19, D14, A1, B20, B22, B24, D28, A3, B18	Fig. 2b
13 channels over the motor cortex and frontal lobe	C29, C17, C16, D11, D13, D19, D27, D17, B32, B30, B22, B19, B17	Fig. 2c
9 channels over the motor cortex	D12, D19, D28, C23, A1, A3, B31, B22, B18	Fig. 2d
9 channels over the motor cortex and frontal lobe	D12, D19, D28, C29, C17, C16, B31, B22, B18	Fig. 2e
6 channels over the motor cortex	D12, D19, D28, B31, B22, B18	Fig. 2f
6 channels over the motor cortex and frontal lobe	C29, C17, C16, D19, A1, B22	Fig. 2g
3 channels over the motor cortex	D19, A1 and B22	Fig. 2h

TABLE II: Parameters of the DeepConvNet Architecture. C is the number of channels and $N = 2$ is number of classes. $T = 3$ s x 128 Hz $= 384$ is the number of samples to consider for classification.

Layer	# filters / (kernel size)	Output
Input		(1, C, $T = 384$)
Conv2D	25 / (1, 5)	(25, C, 380)
Conv2D	25 / (C, 1)	(25, 1, 380)
BatchNorm		(25, 1, 380)
Activation(ELU)		(25, 1, 380)
MaxPool2D	(1, 2)	(25, 1, 190)
Dropout		(25, 1, 190)
Conv2D	50 / (1, 5)	(50, 1, 186)
BatchNorm		(50, 1, 186)
Activation(ELU)		(50, 1, 186)
MaxPool2D	(1, 2)	(50, 1, 93)
Dropout		(50, 1, 93)
Conv2D	100 / (1, 5)	(100, 1, 89)
BatchNorm		(100, 1, 89)
Activation(ELU)		(100, 1, 89)
MaxPool2D	(1, 2)	(100, 1, 44)
Dropout		(100, 1, 44)
Conv2D	200 / (1, 5)	(200, 1, 40)
BatchNorm		(200, 1, 40)
Activation(ELU)		(200, 1, 40)
MaxPool2D	(1, 2)	(200, 1, 20)
Dropout		(200, 1, 20)
Flatten		(4000)
Dense(Softmax)	$N = 2$	(2)

TABLE III: Parameters of the ShallowConvNet Architecture. C is the number of channels and $N = 2$ is number of classes. $T = 3$ s x 128 Hz $= 384$ is the number of samples to consider for classification. The $square$ and log activation functions correspond to x^2 and $log(x)$, respectively.

Layer	# filters / (kernel size)	Output
Input		(1, C, $T = 384$)
Conv2D	40 / (1, 13)	(40, C, 372)
Conv2D	40 / (C, 1)	(40, 1, 372)
BatchNorm		(40, 1, 372)
Activation($square$)		(40, 1, 372)
AveragePool2D	(1, 35), stride (1, 7)	(40, 1, 49)
Activation(log)		(40, 1, 49)
Flatten		(1960)
Dropout		(1960)
Dense(Softmax)	$N = 2$	(2)

TABLE IV: Parameters of the EEGNet architecture. C is the number of channels, T is the number of time samples per channel, F_1 is the number of temporal filters, D is the multiplier depth , F_2 is the number of pointwise filters and N is the number of classes.

We tested set of different parameter (not detailed in the paper). We kept 8 temporal filters $F_1 = 8$, 4 multiplier depth $D = 4$, 16 pointwise filters $F_2 = 16$ and (1, $K = 8$) for the lengths of temporal kernel size for the first convolutional layer.

Layer	# filters / (kernel size)	Output
Input		(1, C, $T = 384$)
Conv2D	8 / (1, K)	(8, C, 384)
BatchNorm		(8, C, 384)
DepthwiseConv2D	4 * 8 / (C, 1)	(32, 1, 384)
BatchNorm		(32, 1, 384)
Activation(ELU)		(32, 1, 384)
AveragePool2D	(1, 4)	(32, 1, 96)
Dropout*		(32, 1, 96)
SeparableConv2D	16 / (1, 16)	(16, 1, 96)
BatchNorm		(16, 1, 96)
Activation(ELU)		(16, 1, 96)
AveragePool2D	(1, 8)	(16, 1, 12)
Dropout*		(16, 1, 12)
Flatten		(192)
Dense(Softmax)	$N = 2$	(2)

TABLE V: Number of parameters for different neural networks

	Number of channels					
Model	128	47	13	9	6	3
DeepConvNet	214,077	163,452	142,202	139,702	137,827	135,952
ShallowConvNet	209,442	79,842	25,442	19,042	14,242	9,442
EEGNet-2.32	3,362	2,066	1,522	1,458	1,410	1,362
EEGNet-4.8	5,794	3,202	2,114	1,986	1,890	1,794